Canadian Food Inspection Agency | Agence canadienne d'inspection des aliments

NGS Applications for Diagnosis and Detection of Plant Pathogens

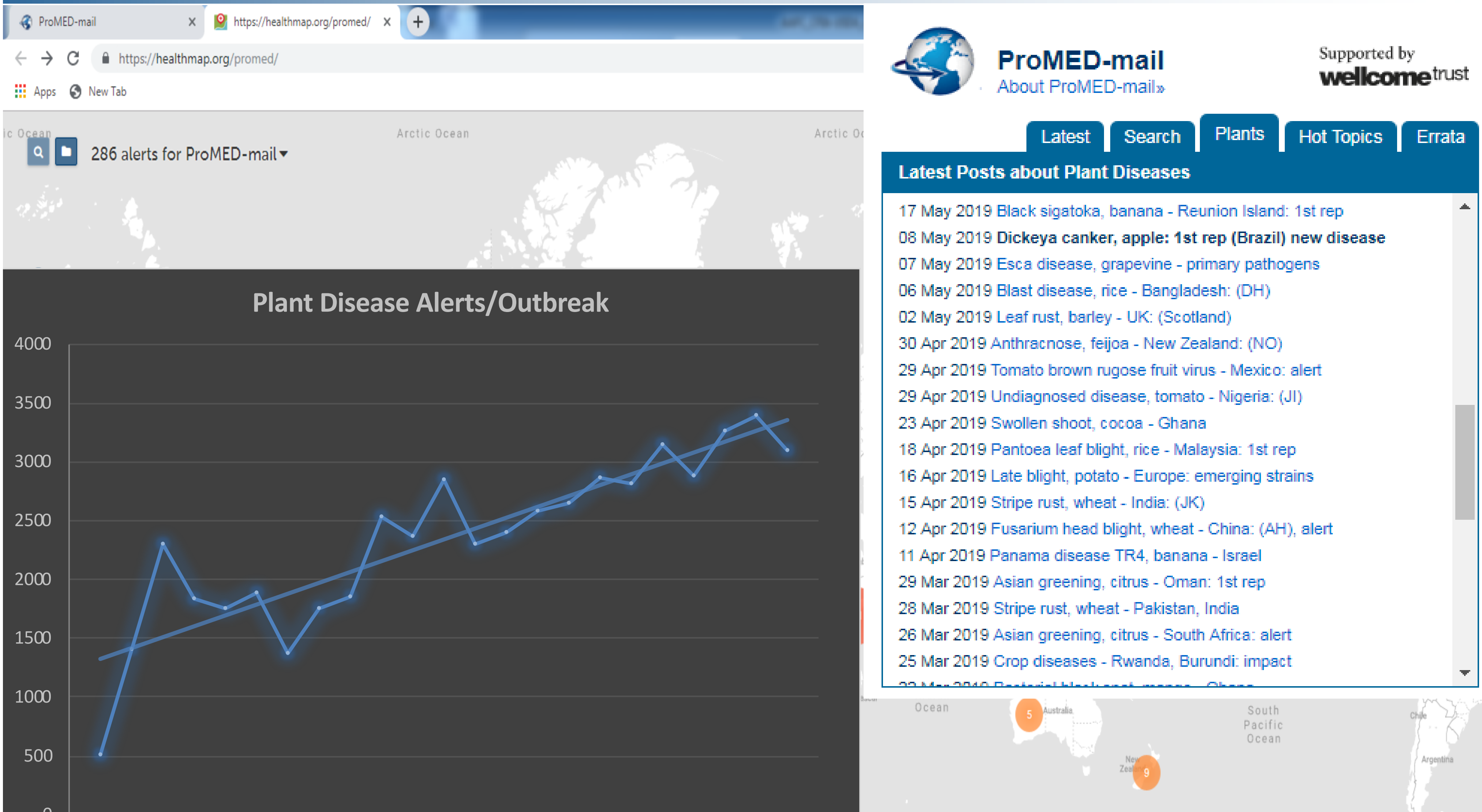
Xiang (Sean) Li, PhD
Canadian Food Inspection Agency
Charlottetown Laboratory
2019.8.3

Canada

NGS Applications for Diagnosis and Detection of Plant Pathogens

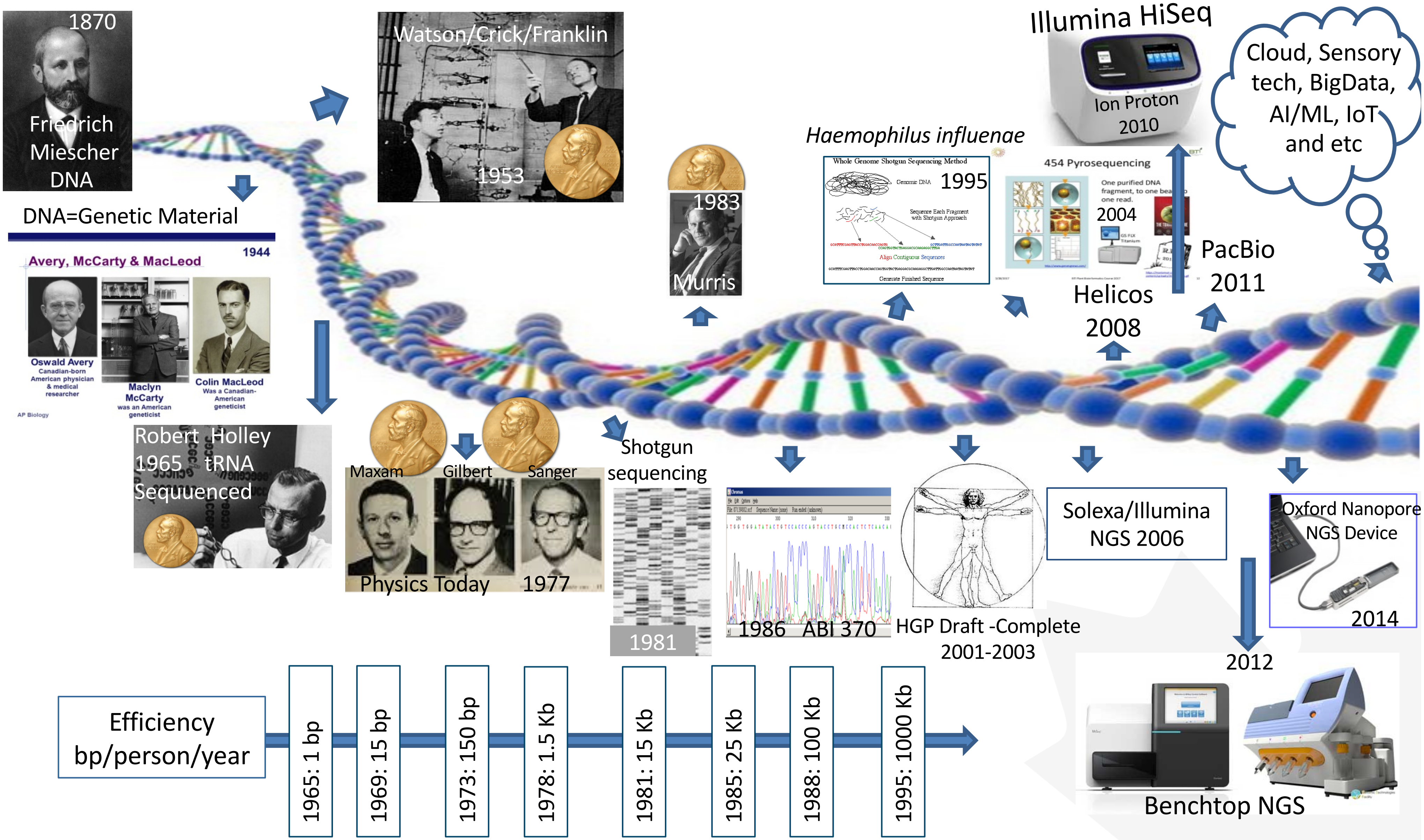
- **Brief Introduction to NGS Technology and Joint Projects**
- **NGS Capability at CFIA (Charlottetown Laboratory)**
- **Simultaneous Detection of Plant Pathogenic Bacterium and Viruses**
- **Summary and Future Prospects**

Current Diagnostics Challenges in Plant Health



http://www.healthmap.org/en/

Brief History on DNA and DNA Sequencing



Big Frame on NGS Software and Databases

- Animal Science Business Line
- Food Science Buisiness Line
- Plant Science Buisiness Line

CFIA CompModel with Multiple Hubs for Bioinformatics and Bigdata Management and Analysis

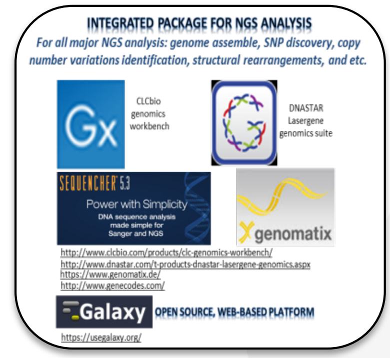
Computation Master Node Charlottetown Laboratory

Internal Storage System

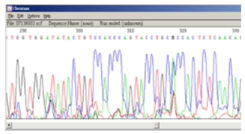
Diagnostics
- Genome workstation

Reference
- Bacterial Refseq database
- Viral Refseq database
- Potato/Tomato Genomes
- Other reference databases

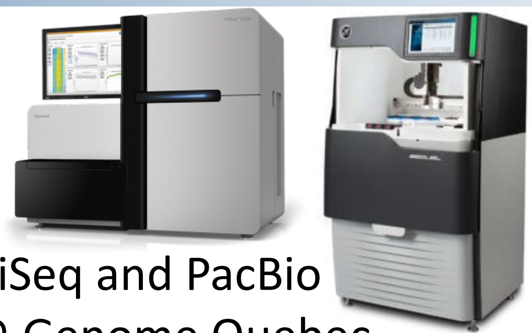
Technolog Development
- Genome workstation
- Genome workstation
- Genome workstation
- Genome workstation



NGS Hardware @ Charlottetown Lab



Ottawa Hospital
Research Institute



HiSeq and PacBio
@ Genome Quebec

NextSeq 500 @ AAFC
Fredericton Lab

Back-up

MiSeq @ Charlottetown Lab

Back-up

Ion PGM @ AAFC
Charottetown Lab

MiSeq & HiSeq
@ BC Cancer Reseach Center

MinION and GridION

PacBio System

NGS Software/Databases @ Charlottetown Lab

Window Based CLC Genomics Workbench

Raw Data

Quality Control

Host Filtration

Alignment Assembly

Downstream Analysis

Computation Master Node @ Charlottetown Laboratory

Internal Storage System

Diagnostics

- Genome workstation

Reference

- Bacterial Refseq database
- Viral Refseq database
- Potato/Tomato Genomes
- Other reference databases

Technology Development

- Genome workstation (Huimin Xu）
- Genome workstation (DL Hammill)*
- Genome workstation (Sean Li)
- Genome workstation (Donna Smith)*
- Genome workstation (Jingbai Nie)*
- Genome workstation (Eric Chuan)*

Linux-Based Bioinformatics Toolkits

AODP 2.4.6.4
Blast2Go
R 3.5.1
Mauve
CytoScape
BWA 0.7.17
BLAST+ 2.7.1
FastQC 0.11.8
AfterQC 0.9.7
SamTools 1.9
BBMap 38.26
BedTools 2.27.1
EDirect 9.90
SPAdes 3.12.0
GenomeTools 1.5.10
TrimGalore 0.4.5
Picard 2.18.14
FastXToolkit 0.0.14
AKATrim an in house trimming tool
ANIGen Toolkit

Applications of NGS @ Charlottetown Lab

Courtesy of Loren Matheson

Virology and PPEQ

Bacteriology

Potato Wart and Nematology

Research Group	Sample Type	Total No	Targets
Huimin Xu DL Hammill	PPEQ	20	**Multiple Viruses with confirmation**
	DIA	10	
	Research	19	
Sean Li JB Nie EJ Chuan	DIA	10	**Bacteria and viruses with confirmation**
	Imported	9	
	Research	15	
D. Smith	Potato wart	40	**Transcriptomes (wart) and artificial constructs (GMO)**
	GMO Trait	7	

RPS Project: CHA-P-1812A

Developing & synchronising methodologies with US toward high throughput detection and diagnostics of bacterial diseases

Project Lead: Dr. Sean Li, CFIA Charlottetown Lab, Charlotetown, Canada

Collaborator: Dr. Michael Stulberg, USDA-APHIS, PPQ, Beltsville, USA

Objectives

1. To evaluate and validate the newly-developed and currently-used qPCR assays at CFIA Charlottetown Laboratory for detecting potato bacterial brown rot caused by *Ralstonia solanacearum* R3bv2 strains with methods employed by USDA-APHIS.
2. To carry out a comprehensive comparison of serological and molecular methods made available during the last 20 years for the detection and identification of *Clavibacter sepedonicus*.

CHA-P-1812A: Developing & Synchronising Methodologies with USDA-APHIS

Project Outcomes

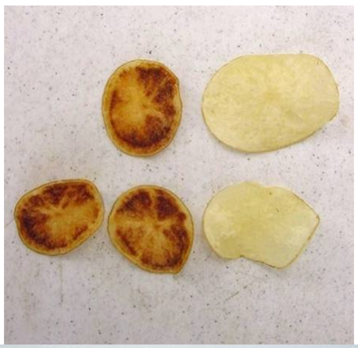
- Generated reliable data for the accurate and sensitive detection of Rs R3bv2 and Cs in potato on the basis of currently available technology.
- Advise the Policy and Programs Branch of new developments in technology and implications for regulatory directives.
- Establish diagnostic schemes with automation component that will best fit in current diagnostic needs, and be transferred to the diagnostic labs of CFIA and other end users, e.g. private labs approved by CFIA.
- Enhance diagnostic capacity at CFIA and gain recognition of Canadian plant pathology expertise and research capacity.
- Publication in scientific journal for relevant research results and discoveries.

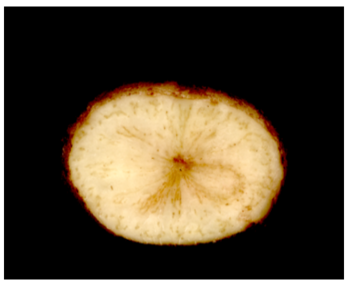
Other Contributors at CFIA with US Collaborators

- **Plant Pathogenic Viruses: Drs. Mike Rott and Huimin Xu are the major contributors at CFIA. A software system, Virtool, has been established for practical applications under evaluation.**
- **Plant Pathogenic Nematodes: CFIA has bilateral connections and discussions with USDA-APHIS regularly.**
- **Potato wart: Dr. Andre Levesque and Donna Smith are the major contributors at CFIA, together with AAFC scientists.**
- **Dickeya-related potato blackleg disease: Dr. Sean Li is the major contributor in CFIA in collaboration with Drs. Amy Charkowski, Colorado State University, and Jianjun Hao, University of Maine.**
- **Bbr caused by *Ralstonia solanacearum* race 3 biovar 2 and Brr caused by *Clavibacter sepedonicus*: Dr. Sean Li is the major contributor in CFIA in collaboration with Drs. Michael Stulberg, USDA-APHIS, and Anne Alvarez, University of Hawaii.**

Simultaneous Detection of PVS and CLso

- Potato zebra chip disease was named after its unique symptoms in potato tubers. It first appeared in Mexico in 1994, and was named in 2008 by Lia W. Liefting and her coworkers as '*Candidatus* Liberibacter solanacearum'.
- The disease was initially observed as darkened bands in processing potato chips, due to an alteration in sugar metabolism.
    - Darkening of the medullary rays ranging from mild to severe in the entire length of the tuber.





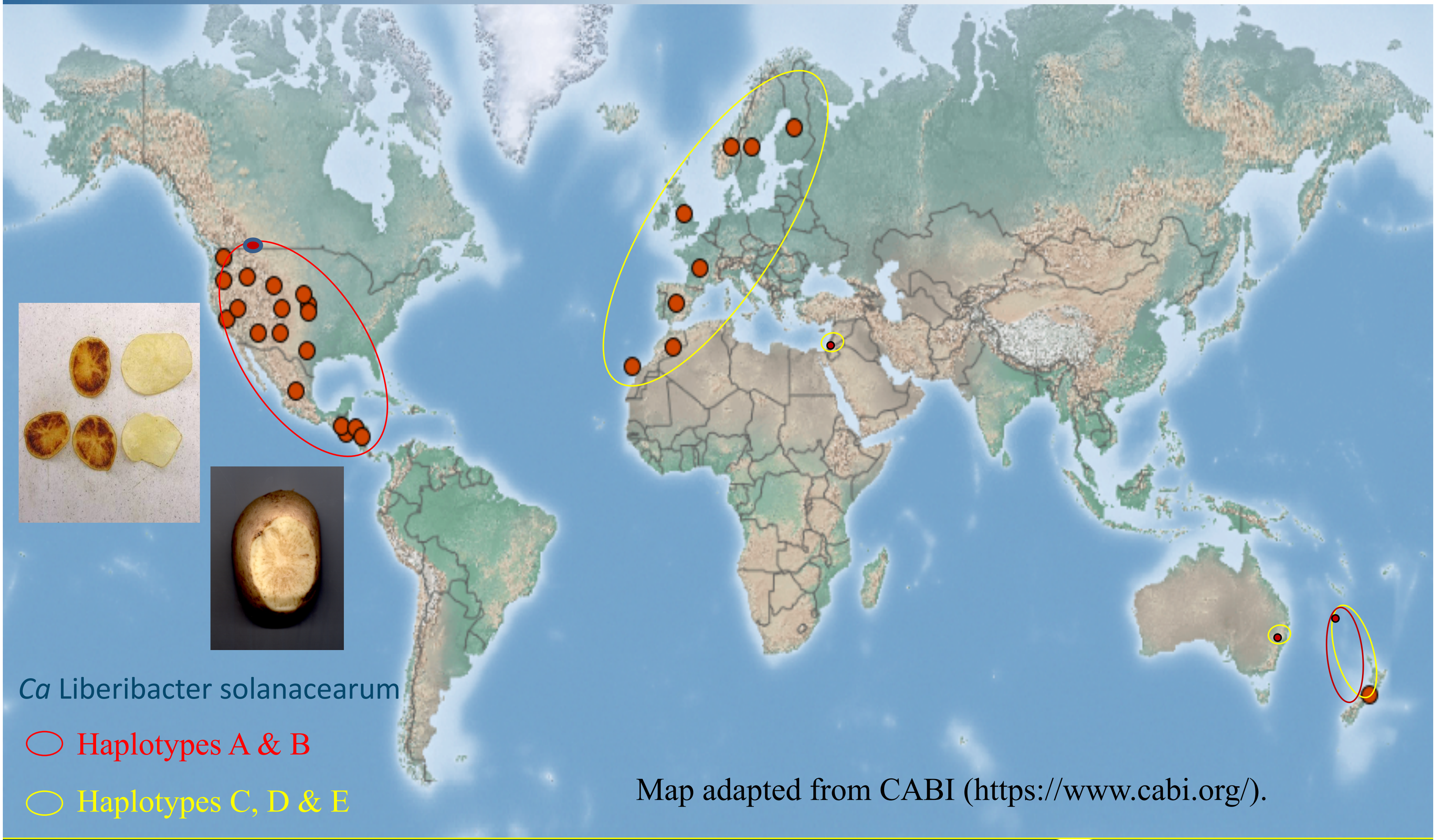
Potato Zebra Chip Disease: the Pathogen and Vector

1. Potato zebra chip disease , also known as "papa manchada" or "papa rayada", is caused by the phloem-limited bacterium '*Candidatus* Liberibacter solanacearum'
2. There are five haplotypes associated with the geographic ranges
   - Haplotypes A&B: Discovered in Mexico, North and Central US, and New Zealand, and vectored by Tomato/potato psyllid;
   - Haplotypes C,D &E: Discovered in Finland and North Europe, and vectored by carrot psyllid.
3. Vectors for the disease transmission include tomato and potato psyllid *Bactericera cockerelli* and carrot psyllid *Trioza apicalis*.

Potato/tomato Psyllids: *Bactericera (Paratrioza) cockerelli*

Carrot psyllid
Trioza apicalis

Global Distribution of Potato Zebra Chip Disease

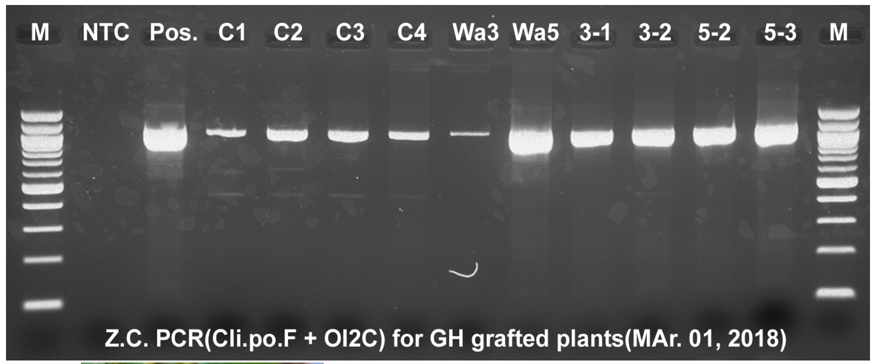


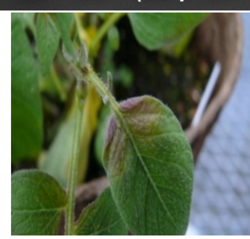
Map adapted from CABI (https://www.cabi.org/).

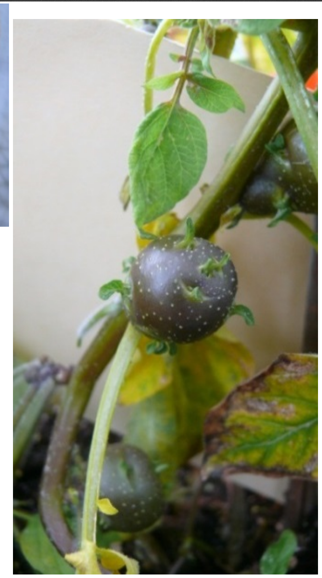
The Virulence of Potato ZC Pathogen

Graft or Psyllid Transmit to Potato or Tomato

- Potato
  - Atlantic, Jemseg
  - Shepody
  - Frontier Russet
  - Russet Norkotah
  - Red Pontiac
- Tomato
  - Roma
  - Moneymaker

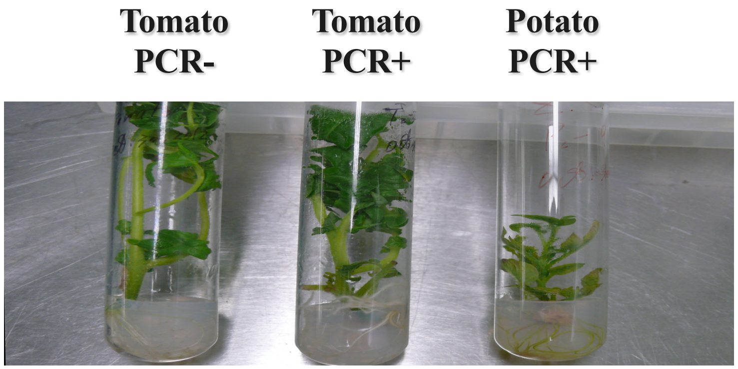


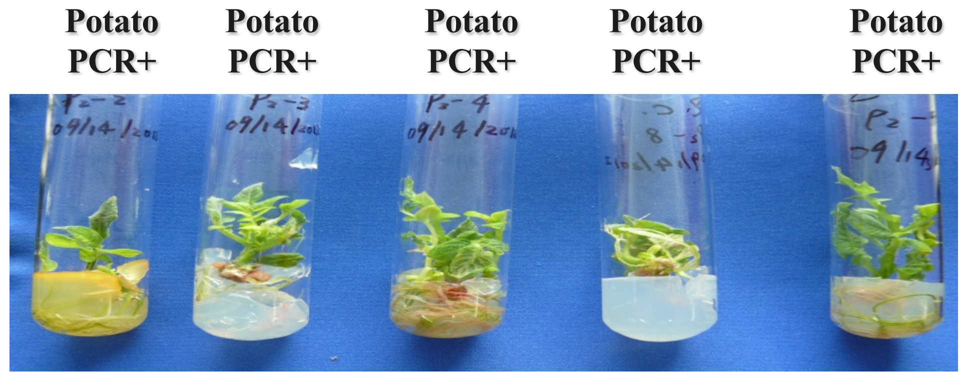




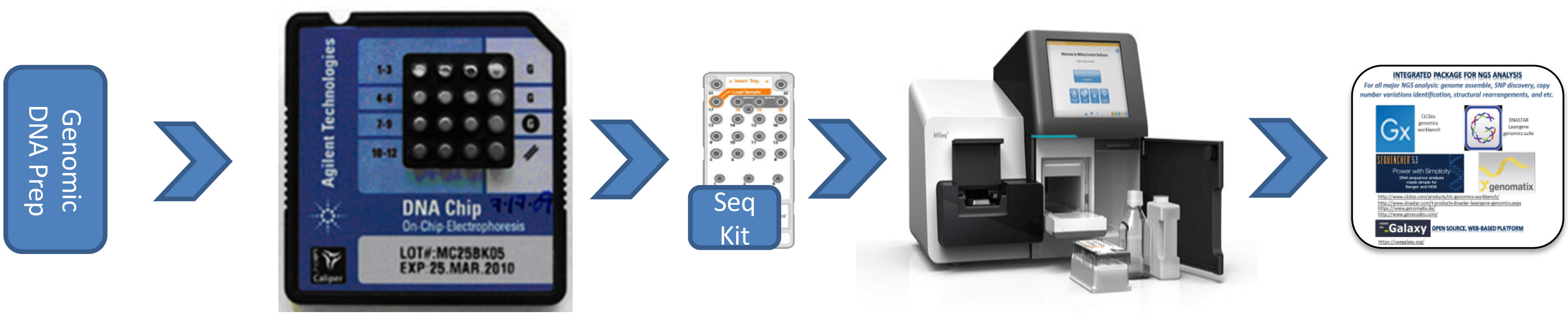
1. Both haplotypes A and B are lethal to potato plants. No resistance or tolerance variety discovered.
2. Haplotype D can be transmitted through carrot seeds, which can be graft-transferred to tomato as well
3. Haplotypes A and D can be preserved in tissue culture tomato plantlets and greenhouse plants.
4. Haplotype B is also lethal to tomato plants, but haplotype A is not lethal to tomato plants.
5. All potato/tomato plants infected by ZC pathogens were tested using PCR followed by sequencing analysis.

Genomics of *Ca* Liberibacter solanacearum

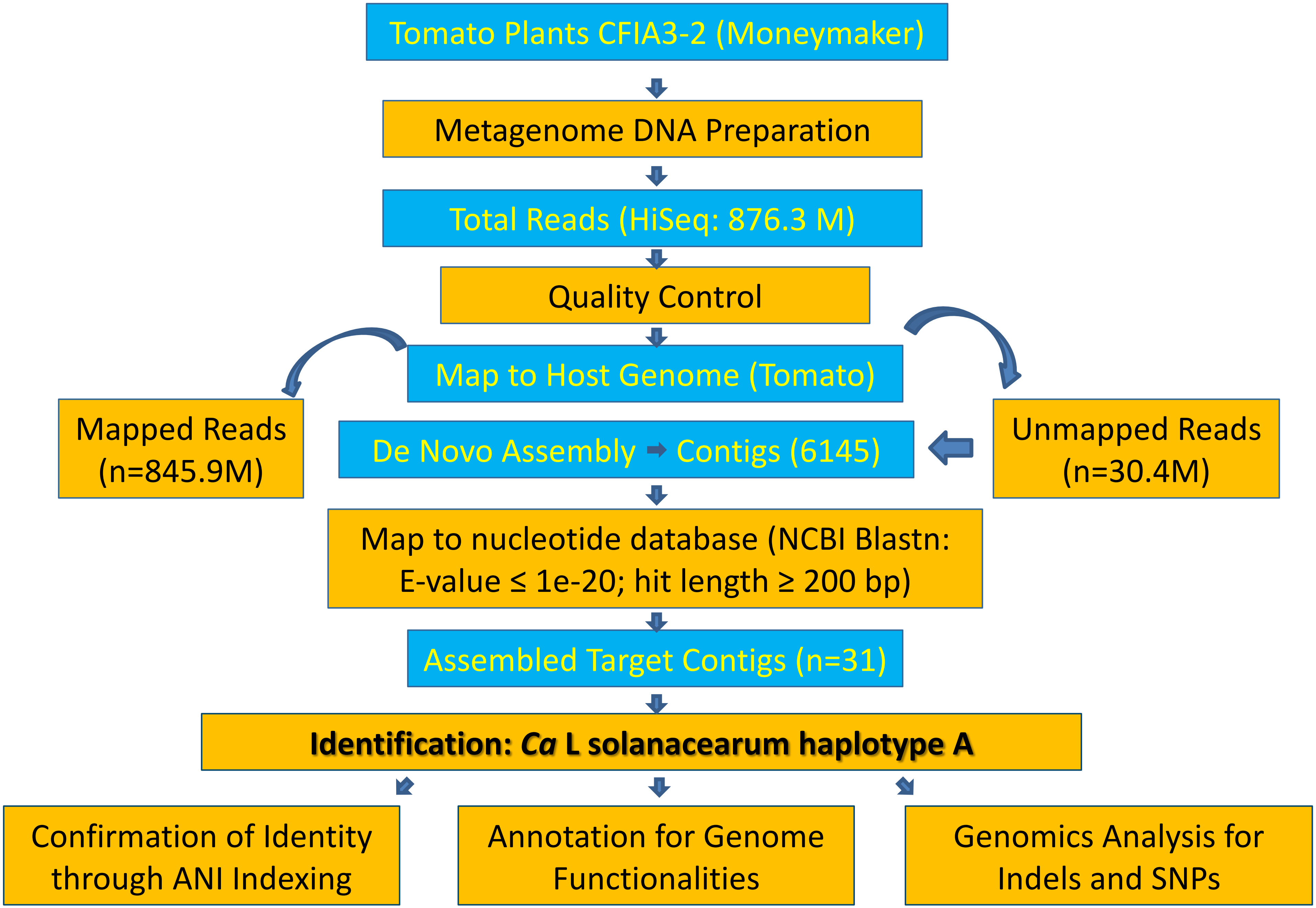




CFIA3-2	The tomato plant infected by haplotype A through grafting nine years ago.	This CLso haplotype A sample originated from potato plants from NDSU. Tomato plants and tissue culture plantlets were graft-transmitted from potato and preserved. All tomato plants graft-transmitted by haplotype B died.
WA5	The tomato plant infected by haplotype A through	The CLso haplotype A sample originated from potato plants from Washington. Tomato plants and tissue culture plantlets were graft



Genomics of *Ca* Liberibacter solanacearum



Simultaneous Detection of PVS and CLso

	CFIA3-2	WA5	C4	Carrot
‘Ca Liberibacter solanacearum’	Haplotype A	Haplotype A	Haplotype D	Haplotype D
Unmapped Read Pairs	510043	1101999	235435	1870698
Mapped Read Pairs	5.72%	11.09%	3.71%	32.55%*
Potato virus S	Negative(PCR-)	Positive(PCR+)	Positive(PCR+)	Negative (PCR-)
Pepino virus	Negative(PCR-)	Negative(PCR-)	Negative(PCR-)	Negative(PCR-)

- **Diagnostic conclusion: CLso haptotype A and D were identified in the four testing plant samples.**
- **Tomato plant WA5 originated from Washington State and C4 originated from Finland contain complete genomes of PVS, whereas tomato sample CFIA3-2 and Carrot sample contain no detectable level of PVS. RT-PCR confirmed the detection of PVS and CLso.**
- **All four samples contained small fragments homologous to Pepino virus (False Positive). PCR failed to detect any trace amount of the virus.**
- **Results were similar using both CLC Workflow and Linux Command Line analyses**

Positive and False Positive

Tomato Plant WA5: Tomato plant grafted with potato with Clso haplotype A NCBI Blastn: E-value ≤ 1e-20; hit length ≥ 80 bp; Bit Score≥80)		
Ca L. solanacearum	Potato Virus S	*Ralstonia solanacearum*
13 Contigs; BScore 179-717	17 Contigs; BScore 286-13K	22 Contigs; BScore ≤ 72
Hit Length: ≥ 218 (713) bp	Hit Length: ≥ 437 (7505) bp	Hit Length: ≤ 61 bp
Homology: ≥ 98%	Homology: ≥ 98%	Homology: ≥ 92%
Positive	Positive	False Positive

Tomato Plant C4: Tomato plant grafted with carrot with Clso haplotype D NCBI Blastn: E-value ≤ 1e-20; hit length ≥ 80 bp; Bit Score≥80)		
Ca L. solanacearum	Potato Virus S	Ralstonia solanacearum
3 Contigs; BScore 425-1379	6 Contigs; BScore 425-1399	13 Contigs; BScore ≤ 62
Hit Length: ≥ 304 (430) bp	Hit Length: ≥443 (7505) bp	Hit Length: ≤ 61 bp
Homology: ≥ 98%	Homology: ≥ 98%	Homology: ≥ 88%
Positive	Positive	False Positive

Metatranscriptomic Analysis (MiSeq vs NextSeq)

Features	CFIA3-2 Haplotype A		WA5 Haplotype A		C4 Haplotype D	
	NextSeq	MiSeq	NextSeq	MiSeq	NextSeq	MiSeq
Read Length	2x150	2x300	2x150	2x300	2x150	2x300
Total Reads	35.4M	17.9M	58.0M	20.1M	144.1M	12.7M
Mapped Reads	34.9M	11.1M	55.7M	17.9M	141.4M	12.4M
Unmapped Reads	526.9K	470.7K	2250.2K	2203.1K	2,752.6K	346.4K
% of Unmapped	1.49%	2.63%	3.88%	10.97%	1.91%	2.72%
Contigs/scaffolds	986	3,617	2344	15,238	1795	980
Blast hits (Clso)	6 (0.6%)	9 (2.5%)	5 (0.2%)	13 (0.09%)	3 (0.17%)	3 (0.3%)
Bit score (Clso)	376-1352	82-539	410-1386	179-717	372-470	425-1399

NextSeq: AAFC Federicton Lab; MiSeq: Charlottetown Lab

Analytic Outcomes of MiSeq Read Length

Tomato Plant B5

Metatranscriptome (RNA)

Total Reads (MiSeq)

Quality Control

Map to Host Genome (Tomato)

Mapped Reads

Unmapped Reads

De Novo Assembly (Contigs)

Map to nucleotide database (NCBI Blastn: E-value ≤ 1e-20; hit length ≥ 80 bp; BScore ≥ 80)

Final Identification: *Ca* L. solanacearum haplotype B

Window Based CLC WB

Metatranscriptome of B5 in MiSeq		Features
2x150	2x300	Read Length
10.6M	11.4M	Total Reads
10.5M	11.1M	Mapped Reads
128.3K	339.9K	Unmapped Reads
1.20%	4.47%	% of Unmapped
167	4,480	Contigs/scaffolds
78 (46.7%)	386 (8.6%)	Blast hits (Clso)
241-1206	212-1886	Bit score (Clso)

Conclusional Remarks and Future Prospects

- Draft genome of Clso haplotype A was obtained through metagenomics approach. Using NGS and bioinformatics approaches, plants infected by bacterial and viral pathogens can be steadily detected simultaneously using NGS Techniques with qPCR confirmations.
- Tomato plant WA5 originated from Washington State and C4 originated from Finland contain complete genomes of PVS, whereas tomato sample CFIA3-2 and Carrot sample contain no detectable level of PVS. All four samples contained small fragments homologous to Pepino virus which were diagnosed as false positives. PCR assay failed to detect any trace amount of the virus.
- Results were confirmed using both CLC Workflow and Linux Command Line analyses Further analysis of differential trascriptomic expression of haplotypes A and B in tomato plants may demonstrate the key mechanisms in virulence and pathogenicity of haplotype B in tomato plants.
- The NGS and bioinformatics workflow are used in assisting diagnostic work on PPEQ and Diagnosis Section at Charlottetown

Acknowledgement

CFIA

Dr. Huimin Xu, Desmond L. Hammill, Jingbai Nie, Kat Yuan, Eric J. Chuan, Donna Smith, Pamela Ross, and many others

Drs. André Lévesque, Mike Rott, Guillaume Bilodeau, Marc-Olivier Duceppe, Burton Blais, Adam Koziol, and many others

AAFC Labs

Drs. André Lévesque, Wen Chen, James Tambong, Iyad Kandalaft,

Xianzhou Nie, Kyle Gardner, Bourlaye Fofana

International Collaborators

Dr. Gary Secor @ NDSU, Dr. Joseph Munyaneza @ USDA-ARS, Dr. Mikko Lehtonen @ Finish Food Safety Authority Eviira,

Dr. Lia Liefting @ Diagnostic and Surveillance Services, Ministry for Primary Industries, New Zealand,

Drs. Charlie Rush, Fedeke Workneh, Li Paetzold @ TAMU

Thank You!